

# SELF CALIBRATION OF GALAXY BIAS IN SPECTROSCOPIC REDSHIFT SURVEYS OF BARYON ACOUSTIC OSCILLATIONS

PENGJIE ZHANG

Shanghai Astronomical Observatory, Chinese Academy of Science, 80 Nandan Road, Shanghai, China, 200030 and  
 Joint Institute for Galaxy and Cosmology (JOINGC) of SHAO and USTC

*Draft version February 18, 2008*

## ABSTRACT

Baryon acoustic oscillation (BAO) is a powerful probe on the expansion of the universe, shedding light on elusive dark energy and gravity at cosmological scales. BAO measurements through biased tracers of the underlying matter density field, as most proposals do, can reach high statistical accuracy. However, possible scale dependence in bias may induce non-negligible systematical errors, especially for the most ambitious spectroscopic surveys proposed. We show that precision spectroscopic redshift information available in these surveys allows for *self calibration* of the galaxy bias and its stochasticity, as function of scale and redshift. Through the effect of redshift distortion, one can simultaneously measure the real space power spectra of galaxies, galaxy-velocity and velocity, respectively. At relevant scales of BAO, galaxy velocity faithfully traces that of the underlying matter. This valuable feature enables a rather model independent way to measure the galaxy bias and its stochasticity by comparing the three power spectra. For the square kilometer array (SKA), this self calibration is statistically accurate to correct for 1% level shift in BAO peak positions induced by bias scale dependence. Furthermore, we find that SKA is able to detect BAO in the velocity power spectrum, opening a new window for BAO cosmology.

*Subject headings:* cosmology: distance scale–large-scale structure of universe–theory

## 1. INTRODUCTION

The universe is the largest laboratory for fundamental physics. The zeroth order descriptions of the universe include the overall expansion rate  $H(z)$  and its integral form, the distance as a function of redshift  $z$ . Type Ia supernovae (SNe Ia), as promising *cosmological standard candles*, allow the measurements of cosmological distance and have enabled the discovery of the late time acceleration of the universe (Riess et al. 1998; Perlmutter et al. 1999). This discovery has profound implication for fundamental physics, leading to either a dominant dark energy component with equation of state  $w \equiv P/\rho < -1/3$  or significant deviations from general relativity around Hubble scale. Ongoing and planned supernova surveys will be able to significantly narrow the range of  $w$  and hopefully clarify the role of the cosmological constant in our universe. Meanwhile, other independent probes are indispensable, to reduce statistical errors, break parameter degeneracies and to cross check for possible systematical errors. The baryon acoustic oscillation (BAO) is such a probe (Blake & Glazebrook 2003; Hu & Haiman 2003; Seo & Eisenstein 2003, 2005; Blake et al. 2006).

Prior to the epoch of recombination, baryons and photons are tightly coupled. Fluctuations in the baryon-photon fluid propagate as acoustic waves at speed comparable to the speed of light. These acoustic waves imprint in the cosmic microwave background, a snapshot of the photon fluid at the epoch of decoupling. They also imprint in the baryon fluid through coupling between photons and baryons and later in the overall matter density field, through gravitational coupling between baryons and dark matter particles (Hu & Sugiyama 1996; Eisenstein & Hu 1998). These BAO features have been detected and measured to high precision in CMB and have put strong constraints on the geometry of the

universe (Spergel et al. 2007). In the large scale structure (LSS) of the universe, these features are significantly suppressed by dark matter to a level much weaker than that in CMB. Nonetheless, they are detectable, as shown in measurements of correlation function and power spectrum of 2dF and SDSS galaxies (Cole et al. 2005; Eisenstein et al. 2005; Percival et al. 2007). The comoving scales of BAOs can be predicted accurately from well understood linear physics. Thus we are able to convert the measured angular separations of these features on the sky to the cosmological angular diameter distance, free of many astrophysical uncertainties. Furthermore, since we can measure their separations in redshift, we are able to directly measure  $H(z)$ , which is more sensitive to  $w$  than the distance. For these merits, BAO is widely accepted as a robust and powerful *cosmological standard ruler* to reveal the expansion history of the universe (Albrecht et al. 2006).

All existing proposals on BAO measurements target at density field of various tracers of the LSS. Most are galaxy spectroscopic redshift surveys, including LAMOST<sup>1</sup>, WiggleZ (Glazebrook et al. 2007), BOSS<sup>2</sup>, WFMOS (Bassett et al. 2005), HETDEX<sup>3</sup>, ADEPT<sup>4</sup> and the square kilometer array (SKA, Abdalla & Rawlings 2005). Besides, there are proposals to measure BAO in the 21cm background, both at low redshifts (Chang et al. 2007) and at high redshifts (Mao & Wu 2007; Wyithe et al. 2007), in the Lyman-alpha forest (McDonald & Eisenstein 2007) and in SNe Ia (Zhan et al. 2008). Sub-percent level statistical accuracy can be reached in the most ambitious surveys,

<sup>1</sup> <http://www.lamost.org/en/>

<sup>2</sup> <http://cosmology.lbl.gov/BOSS/>

<sup>3</sup> <http://www.as.utexas.edu/hetdex/>

<sup>4</sup> [http://www7.nationalacademies.org/ssb/BE\\_Nov\\_2006\\_bennett.pdf](http://www7.nationalacademies.org/ssb/BE_Nov_2006_bennett.pdf)

such as SKA, classified as a stage IV project by the dark energy task force (Albrecht et al. 2006). However, possible sources of systematical errors could prohibit us to exploit full advantage of these measurements. Some, for example those induced by gravitational lensing (Vallinotto et al. 2007; Hui et al. 2007a,b; LoVerde et al. 2007), are correctable (Zhang & Chen 2007; Hui et al. 2007b). Nonlinear evolution in the matter density field is also likely correctable, either by improvement in theoretical calculation (Croce & Scoccimarro 2006, 2007; Matarrese & Pietroni 2007a,b; McDonald 2007) or in data analysis (Eisenstein et al. 2007b). However, the clustering bias of these tracers with respect to the underlying matter density field, is more elusive to handle. The scale dependence of the bias changes the overall shape of the power spectrum and thus shifts the BAO positions. A good thing is that, to the first order accuracy, BAOs, as periodic sharp features on top of the smooth power spectrum, are robust against bias scale dependence, which is likely more smooth. For this reason and others, systematical errors caused by possible scale dependent bias are likely sub-dominant, comparing to the statistical errors, for BAO measurements of existing surveys and proposed medium cost surveys. On the other hand, the most ambitious surveys such as SKA will have about two orders of magnitude increase in survey volume and one order of magnitude increase in redshift over SDSS. For these surveys, statistical errors in the distance and  $H(z)$  measurement will be reduced to below 1% (Blake et al. 2006). At this stage, shifts in BAO peaks induced by even rather mild scale dependence in galaxy bias may become non-negligible. For discussions on the effect of bias to BAO, refer to e.g. Seo & Eisenstein (2005); Eisenstein et al. (2007a); Guzik et al. (2007); Huff et al. (2007); Smith et al. (2007a); Angulo et al. (2008).

Fortunately, spectroscopic redshift information, which is required to measure  $H(z)$  along the line of sight through BAO, allows for self calibration of galaxy bias through the effect of redshift distortion. Spectroscopic redshift surveys are able to measure  $P_g^s(\mathbf{k})$ , the 3D power spectrum of galaxies in redshift space, to high precision. Galaxy peculiar velocities make  $P_g^s$  anisotropic in the 3D wavevector  $\mathbf{k}$  space. This redshift distortion effect is often treated as a source of error. However, it can be rendered into valuable source of signal instead, given sufficiently accurate understanding of redshift distortion.

Through the anisotropy in redshift space, one can reconstruct the real space velocity power spectrum  $P_\theta$ , the velocity-galaxy cross correlation power spectrum  $P_{g\theta}$  and the galaxy power spectrum  $P_g$  in a rather model independent way at cosmological distances. Here  $\theta$  is the normalized divergence of peculiar velocity. Since galaxies have rather low mass, they are sub-dominant in gravitational field at scales  $\gtrsim 10\text{Mpc}$ . At scales of BAO (much larger than 10 Mpc, even for the third peak), gravity is the only force accelerating galaxies. For these reasons, galaxies can be well approximated as test particles, following the motion of dark matter particles. Thus the large scale galaxy velocity field is a faithful tracer of the overall matter velocity field.  $P_\theta$  and its relation with the matter power spectrum  $P_m$  can then be predicted from first principles, free of uncertainties in understanding of galaxy formation. Comparing  $P_g$ ,  $P_{g\theta}$  and  $P_\theta$ , we are

able to measure galaxy bias and its stochasticity. This allows for self calibration of galaxy bias in BAO cosmology. In a pioneer work, Pen (1998) proposed to measure galaxy bias and its stochasticity through the quadrupole and octupole of  $P_g^s$ . This work extends his by a full power spectrum reconstruction, for which we are able to utilize information contained in all moments.

On the other hand, BAOs also exist in the velocity field.  $P_\theta$  reconstruction thus opens a new window of BAO measurement. This approach is free of problems of galaxy bias, although it may not be able to reach the same statistical accuracy as that of the density field.

This technique can be incorporated into existing BAO analysis methods, such as that of Seo & Eisenstein (2003, 2007) and Blake & Glazebrook (2003); Blake et al. (2006); Percival et al. (2007), as straightforward post processing. We will explain this issue later.

SKA is able to detect billions of galaxies over half sky to redshift  $z \gtrsim 3$  through neutral hydrogen 21cm emission of these galaxies and measure precise spectroscopic redshifts without extra cost. For this reason, we choose SKA as our primary target for investigation. We present the reconstruction procedure in §2, demonstrate the feasibility of SKA to self calibrate galaxy bias in §3, show that BAOs in the galaxy velocity field are detectable through SKA in §4 and discuss in §5.

## 2. VELOCITY RECONSTRUCTION THROUGH REDSHIFT DISTORTION

Peculiar motions of galaxies imprint unique signatures in the redshift space galaxy power spectrum  $P_g^s(\mathbf{k})$ , which takes a general form (e.g. Scoccimarro (2004)),

$$P_g^s(\mathbf{k}) = [P_g(k) + 2u^2 P_{g\theta}(k) + u^4 P_\theta(k)] F\left(\frac{k^2 u^2 \sigma_v^2}{H^2(z)}\right) \quad (1)$$

Here,  $P_g$ ,  $P_{g\theta}$ ,  $P_\theta$  are the real space power spectra of galaxies, galaxy-velocity and velocity.  $\theta$  is the comoving peculiar velocity divergence divided by  $-H(z)$ .  $\sigma_v$  is the 1D velocity dispersion; and  $F(x)$  is a smoothing function, normalized to unity at  $x = 0$ , determined by the velocity probability distribution.  $u = k_\parallel/k$  is the cosine of the angle of the  $\mathbf{k}$  vector with respect to radial direction; This unique directional dependence has enabled successful simultaneous reconstruction of the three power spectra from 2dF and SDSS galaxies (Tegmark et al. 2002, 2004). We adopt a simple minimum variance estimator, developed in Zhang et al. (2007), to quantify accuracies of this approach.

For each  $\mathbf{k}_i$  in the given  $k$  bin, we have a measurement of  $P_g^s$ , which we denote as  $P_i$ . The unbiased minimum variance estimator of the band power of  $P_{(\alpha)}$  ( $\alpha = g, g\theta, \theta$ ), is  $\hat{P} = \sum W_i^{(\alpha)} P_i$ , where  $W_i^{(\alpha)} = \frac{F_i}{2\sigma_i^2} (\lambda_1^{(\alpha)} + \lambda_2^{(\alpha)} u_i^2 + \lambda_3^{(\alpha)} u_i^4)$  and  $F_i \equiv F(k u_i \sigma_v / H)$ .  $\sigma_i = P_i + 1/n_g$  is the rms fluctuation of  $P_i$ , where  $n_g$  is the galaxy number density. The Lagrange multipliers  $\lambda_{1,2,3}^{(\alpha)}$  are given by  $\lambda^{(g)} = (1, 0, 0) \cdot \mathbf{A}^{-1}$ ,  $\lambda^{(g\theta)} = (0, 1/2, 0) \cdot \mathbf{A}^{-1}$  and  $\lambda^{(\theta)} = (0, 0, 1) \cdot \mathbf{A}^{-1}$ .  $\lambda^{(\alpha)}$  are orthogonal to each other. The  $3 \times 3$  matrix  $\mathbf{A}$  is given by

$$A_{mn} = \sum_i u_i^{2(m+n-2)} \frac{F_i^2}{2\sigma_i^2} ; m, n = 1, 2, 3. \quad (2)$$

The corresponding error in each power spectrum is

$$\sigma_p^{(\alpha)} = \left( \frac{1}{2} \lambda^{(\alpha)} \cdot \mathbf{A} \cdot [\lambda^{(\alpha)}]^T \right)^{1/2}. \quad (3)$$

One can check that it has the right scaling that  $\sigma_p \propto \sigma/\sqrt{N}$ , where  $N$  is the number of independent modes and  $\sigma$  is some average of  $\sigma_i$ . This relation can be further simplified and we have  $\sigma_p^{(g)} = \sqrt{\lambda_1^{(g)}/2}$ ,  $\sigma_p^{(g\theta)} = \sqrt{\lambda_2^{(g\theta)}/4}$  and  $\sigma_p^{(\theta)} = \sqrt{\lambda_3^{(\theta)}/2}$ .

Combining these power spectrum measurements, we are able to measure the galaxy bias and its stochasticity through

$$b_g^2 = \left( \frac{P_g}{P_\theta} \right) \left( \frac{P_\theta}{P_m} \right)_{theory}; \quad r = \frac{P_{g\theta}}{\sqrt{P_g P_\theta}}. \quad (4)$$

Here,  $(P_\theta/P_m)_{theory}$  is predicted from theory. In linear regime, this value is  $f^{-2}$ , where  $f \equiv (d \ln D / d \ln a)^2$  and  $D$  is the linear density growth factor.  $r$  is the cross correlation coefficient, whose deviation from unity is a measure of stochasticity (Pen 1998; Dekel & Lahav 1999).

This reconstruction requires the angular diameter distance  $D_A$  and  $H(z)$  as input to convert the observed angular and radial separation into  $\mathbf{k}$ . For numerical results presented in this paper, we just take  $D_A$  and  $H(z)$  of the fiducial cosmology as input. For real data, a convenient and self consistent procedure to carry out this reconstruction is as follows. First, one can measure the distance and  $H(z)$  by the usual methods (Blake & Glazebrook 2003; Seo & Eisenstein 2003). The measured distance and  $H(z)$  are then used as input for the reconstruction of the power spectra, which are in turn applied to check for the consistency of scale independent bias. This can be done iteratively. Alternatively, one can allow for scale dependent bias and its stochasticity in BAO analysis method of Seo & Eisenstein (2003, 2007). Through simultaneous multi-parameter fitting, galaxy bias (amplitude, scale dependence and stochasticity) is automatically self calibrated.

Errors in  $D_A$  and  $H(z)$  propagate into the power spectrum reconstruction, through two effects: a constant fractional shift in  $k$  and a wrong determination in the  $\mathbf{k}$  direction  $u$ . Obviously, the overall shift in  $k$  does not affect the measurement of  $b_g$  and  $r$ . However, a wrong determination in  $u$  does.  $P_{g\theta}$  and  $P_\theta$  are determined by the derivatives of  $P_g^s(u)$  with respect to  $u$ . We can show that the induced errors in the power spectra have only weak dependence on  $k$  through the slope of the power spectra. Thus the main effect of errors in  $D_A$  and  $H(z)$  is to bias the overall amplitude of  $b_g$  and  $r$ , instead of introducing new scale dependence. For these reasons, it suffices to neglect this kind of error source.

An implicit simplification adopted in this reconstruction is that  $\sigma_v$  is given. In reality, at BAO relevant scales, the function  $F(ku\sigma_v/H) \simeq 1$ . Thus our simplification is justified here. For real data, to improve the accuracy of reconstruction, one should treat  $\sigma_v$  as a free parameter to be marginalized over. We do not expect any major effect on the forecasted errors in the reconstructed power spectra.

Hereafter we will adopt the linear theory to carry out numerical calculations. For real data, nonlinearity must

be taken into account in BAO analysis. For example, nonlinear evolution in density field and velocity field are not the same (Bernardeau et al. 2002; Smith et al. 2007b). As a result,  $(P_\theta/P_g)_{theory}$  is no longer equal to  $f^{-2}$  and is no longer scale independent. If unaccounted, we may obtain a false scale dependence in  $b_g$ . However, this should not be a real problem, due to three reasons. (1) The reconstruction does not rely on assumptions of linearity. (2) The bias is defined with respect to the real (thus nonlinear) power spectra. (3) Nonlinear evolution in the velocity field and matter field can be calculated from first principles. As long as we replace the linear theory version  $f^{-2}$  with the actual (nonlinear)  $(P_\theta/P_m)_{theory}$  in real data analysis, nonlinearity does not bias the result.

### 3. SELF CALIBRATION OF GALAXY BIAS

In this section, we choose SKA as the primary target for investigation. For the redshift range  $1 \lesssim z \lesssim 2$  of most interest, the SKA survey speed  $(A_{\text{eff}}/T_{\text{sys}})^2 \times \text{FOV} = 2 \times 10^{10} \text{ m}^2 \text{ K}^{-2} \text{ deg}^2$ .<sup>5</sup> With this specification, SKA is able to detect several billions of 21cm emitting galaxies in 10 year survey. Throughout this paper, we adopt the galaxy surface number density as 20 galaxies per square arc-minute, corresponding to 2.4 billion over three quarter of sky. The exact number is hard to estimate, due to poor knowledge on these galaxies at high redshifts. A good thing is that the exact number is not necessary for the estimation presented in this paper, due to the dominance of cosmic variance over shot noise, especially at scales larger than the second baryon acoustic peaks.

To proceed, we adopt the fiducial cosmology as a flat  $\Lambda$ CDM cosmology, with  $\Omega_m = 0.268$ ,  $\Omega_b = 0.044$ ,  $\Omega_\Lambda = 1 - \Omega_m$ ,  $h = 0.72$  and  $\sigma_8 = 0.77$  (Spergel et al. 2007). We adopt the transfer function fitting formula of Eisenstein & Hu (1998). We stick to the linear perturbation theory and neglect the nonlinear evolution in the matter density field, since it is predictable and correctable. The velocity power spectrum is fixed through  $P_\theta = f^2 P_m$ . Here,  $f = d \ln D / d \ln a$ ,  $D$  is the linear density growth factor and  $P_m$  is the matter power spectrum.

To demonstrate the feasibility of self calibration, we introduce scale dependence and stochasticity in galaxy distribution. As a reminder, the galaxy bias  $b_g(k, z)$  is defined by  $b_g^2(k, z) \equiv P_g(k, z)/P_m(k, z)$ . The stochasticity of galaxy distribution is characterized by the cross correlation coefficient  $r(k, z)$ , defined through  $r^2 \equiv P_{g\theta}^2/(P_g P_\theta)$ . Deterministic bias will have  $r = 1$ . This is likely true at very large scales. Proceeding to smaller scales we may expect  $r < 1$ , because galaxy formation depends on not only the local density, but also the environment. We adopt simple forms of parametrization:  $b_g = b_{g0} + db_g/dk|_{k=0}k \equiv b_{g0}(1 - k/k_b)$  and  $r = 1 + dr/dk|_{k=0}k \equiv 1 - k/k_c$ . Here,  $b_{g0}$  is the galaxy bias when  $k \rightarrow 0$ . We further adopt  $k_b = 4h/\text{Mpc}$  and  $k_r = 10h/\text{Mpc}$ . The actual scale dependence of bias can be much more complicated, with terms  $\propto k$ ,  $\propto k^2$ ,  $\propto 1/P_m(k)$ , or even more complicated terms (refer to, e.g. Smith et al. 2007a). We choose the above simple

<sup>5</sup> [http://www.skatelescope.org/pages/Preliminary\\_specifications\\_v2.4.pdf](http://www.skatelescope.org/pages/Preliminary_specifications_v2.4.pdf).

This survey speed is much higher than what adopted in previous estimations of Abdalla & Rawlings (2005); Zhang & Pen (2006), resulting in much higher sensitivity to detect high redshift galaxies.

forms of parametrization and values for several reasons. (1) These forms of parametrization are just the Taylor expansion of  $b_g$  and  $r$  at  $k = 0$  to first order. Despite their simplicity, they are rather general to describe small deviations from scale independence. (2) For the values adopted, the galaxy bias decreases by less than 5% over the first three BAO peaks. This scale dependence has the same sign as, but smaller amplitude than, what predicted for low mass galaxies (Sheth & Tormen 1999; Guzik et al. 2007; Smith et al. 2007a). Thus the assumed scale dependence likely represents a conservative lower limit of this source of systematical errors. (3) Such level of scale dependence in  $b_g$  may cause non-negligible systematical errors. It shifts the positions of the second and the third peak by  $\sim 1\%$ . This level of systematic shift in BAO positions could overwhelm the statistical uncertainties in SKA BAO measurements.<sup>6</sup> For these reasons, the assumed bias scale dependence and stochasticity are suitable to demonstrate the feasibility of the self calibration technique.

We remind the readers that, these parametrizations of  $b_g(k, z)$  and  $r(k, z)$  only serve as input of the fiducial model. The reconstruction of  $P_g$ ,  $P_{g\theta}$  and  $P_\theta$  does not rely on any assumptions on  $b_g$  and  $r$ . Parametrizations on  $b_g$  and  $r$  can certainly improve over the statistical errors in the reconstruction. However, inappropriate parametrizations could bias the reconstruction. For this reason, we do not explore such approach in this paper.

For SKA,  $P_g$  can be reconstructed to high precision (Fig. 1) and the associated fractional statistical errors are very close to that of corresponding  $P_g^s$ . Accuracies in  $P_{g\theta}$  reconstruction are about a factor of 5 poorer, but still impressive. A further factor of  $\sim 3$  degradation occurs in  $P_\theta$  reconstruction. This is what we expected. The average contribution of  $P_\theta$  to  $P_g^s$  is  $P_\theta \langle u^4 \rangle = P_\theta/5 \simeq P_m \Omega_m^{1.2}(z)/5$  and that of  $P_{g\theta}$  is  $2P_{g\theta} \langle u^2 \rangle \simeq P_m 2\Omega_m^{0.6}(z)/3$ . The realistic galaxy bias is likely  $b_g \gtrsim 1/2$  (Fig. 2 and further reading in Jing 1998). We thus have  $P_\theta \langle u^4 \rangle < 2P_{g\theta} \langle u^2 \rangle < P_g$ . Furthermore, we rely on the directional dependence of  $P_g^s$  to extract sub-dominant components  $P_{g\theta}$  and  $P_\theta$ , resulting in  $\sigma_p^{(\theta)}/P_\theta > \sigma_p^{(g\theta)}/P_{g\theta} > \sigma_p^{(g)}/P_g$ .

Nonetheless, the reconstructed power spectra (especially  $P_\theta$ ) have the statistical accuracy to detect 5% or even smaller variation in the galaxy bias over the first three peaks, with large statistical significance (Fig. 1). This 5% scale dependence causes 1% shift in BAO peak positions. Thus the self calibration technique is able to reduce systematical error in BAO peak positions induced by scale dependent bias to smaller than 1%.

However, despite its promising capability, the self cal-

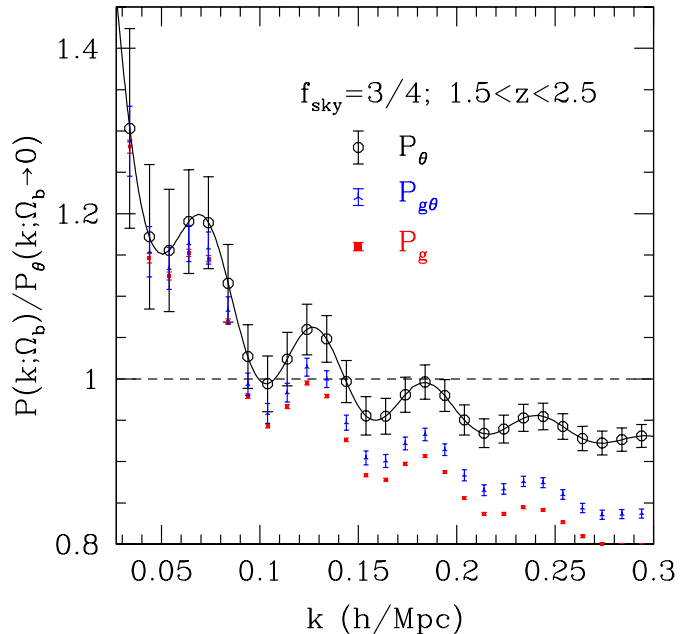


FIG. 1.— Forecasted accuracies in reconstructed power spectrum  $P_g$ ,  $P_{g\theta}$  and  $P_\theta$ . The estimation is based on SKA ten year survey over three quarters of sky and galaxies at  $1.5 < z < 2.5$ . We plot the ratio of three power spectra with respect to  $P_\theta(\Omega_b \rightarrow 0)$  and normalized the ratios to be equal at  $k \rightarrow 0$ . Error bars are forecasted adopting  $b_{g0} = 1$ .  $P_\theta$  is much harder to measure than  $P_g$  and  $P_{g\theta}$ , so the self calibration is fundamentally limited by the associated error in  $P_\theta$ . Nonetheless, SKA is able to test and correct for several percent level scale dependence in bias across the first three peaks. Furthermore, with such accuracy, BAOs are detectable in  $P_\theta$ . Accuracy in  $P_\theta$  is sensitive to  $b_g$ , as shown in Fig. 3.

ibration procedure is fundamentally limited by accuracy in our understanding of redshift distortion. Systematic deviation in the modelled  $P_g^s$  from the actual one can propagate into the estimation of  $b_g$  and thus can bias the self calibration of galaxy bias. A good news is that, to measure the scale dependence and perform self calibration, we are not necessarily confined to BAO scales. As seen from Fig. 1, we are able to reconstruct the three power spectra to  $k \sim 0.3h/\text{Mpc}$  and even smaller scales.  $k = 0.3h/\text{Mpc}$  is still in the linear regime at  $z = 2$ . This means that we have a factor of  $\sim 2$  larger (and still safe) dynamical range to measure the scale dependence in  $b_g$ . Furthermore, variation in  $b_g$  over this larger dynamical range is likely larger. Both alleviate the accuracy requirement for modeling the redshift distortion (e.g. Eq. 1). For the simple model we adopted, this means to measure 8% variation in  $b_g$  over  $k \lesssim 0.3h/\text{Mpc}$  instead of measure  $\sim 5\%$  variation over  $k < 0.2h/\text{Mpc}$ . Still, the modeling of the redshift space power spectrum must be accurate to much better than  $2 \times 8\% = 16\%$  percent in order to render the associated systematical error negligible. This requirement is challenging, but not impossible, given the rapid development in simulations and analytical calculations. We refer the readers to works on modeling of redshift distortion, through perturbation theory (Scoccimarro 2004; Matsubara 2007), halo model (White 2001; Seljak 2001; Tinker 2007) and N-body simulations

<sup>6</sup> The exact level of systematical errors induced by scale dependent bias varies with BAO analysis methods. The method of Seo & Eisenstein (2003) performs the global fitting to the galaxy power spectrum and utilizes the cosmological information in the overall shape of the power spectrum. On the other hand, the method of Blake & Glazebrook (2003); Percival et al. (2007) discards such information. It fits the overall shape of  $P_g$  with a smooth reference function  $P_{\text{ref}}$ . It then fits  $P_g/P_{\text{ref}}$  by a decaying sinusoidal function, in which the BAO information is encoded.  $P_{\text{ref}}$  can capture and thus filter away some of, if not most of, the bias scale dependence. Thus this method is less affected by the scale dependence in galaxy bias. But the residual could still bias the distance measurement.

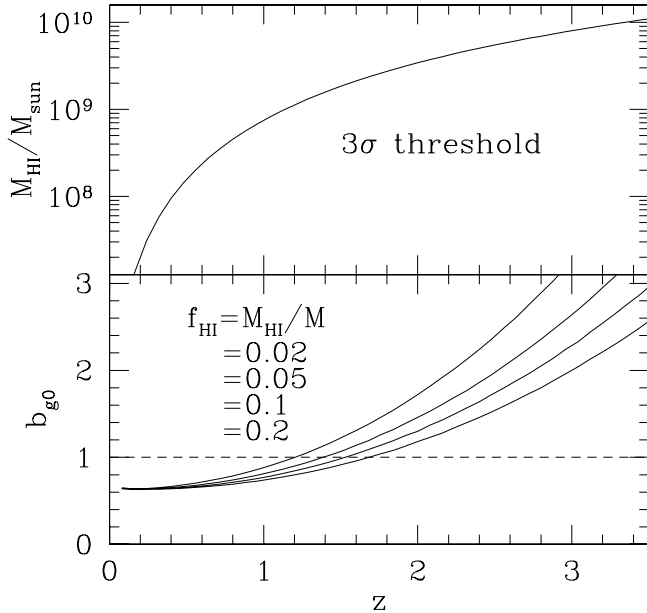


FIG. 2.— The SKA galaxy selection threshold (top panel) and the corresponding bias (bottom panel). SKA galaxies are selected by the galaxy neutral hydrogen mass  $M_{\text{HI}}$ , while the bias is determined by the total galaxy mass  $M$ . Lack of robust measurements on  $f_{\text{HI}} \equiv M_{\text{HI}}/M$ , we plot the galaxy bias for several possible values of  $f_{\text{HI}}$ .

(Kang et al. 2002; Tinker et al. 2006).

#### 4. DETECTING BAO IN THE VELOCITY POWER SPECTRUM

BAOs in the density field inevitably show up in the gravitational potential field, as directly seen from the Poisson equation. This featured gravitational field accelerates galaxies and thus imprints BAOs in the velocity power spectrum  $P_\theta$ . At scales of BAOs, the galaxy velocity field should be unbiased with respect to the velocity field of the underlying matter, since these galaxies are test particles responding only to gravity at these scales. For this reason, we are able to calculate  $P_\theta$  from first principles. Since in principle there is no ambiguity in precision calculation of  $P_\theta$ , BAO cosmology based on  $P_\theta$  is free of the issue of galaxy bias. It is thus worthy of detecting BAOs in  $P_\theta$ .

In reality, BAO measurements through reconstructed  $P_\theta$  from redshift distortion are challenging, since accuracies in  $P_\theta$  are much worse than that in  $P_g^s$  or  $P_g$ , as explained in last section. For  $b_{g0} = 1$ , the fractional error in  $P_\theta$  is a factor of  $\sim 10$  higher than that in  $P_g$  (Fig. 1). It thus requires  $\sim 100$  times increase in survey volume in order to reach the same statistical accuracy obtained from the density field. For this reason, BAO can not be detected in  $P_\theta$  by those low redshift spectroscopic surveys such as LAMOST and BOSS. But SKA can (Fig. 1 & 3), so as the Hubble sphere hydrogen survey (HSBS) (Peterson et al. 2006) and ADEPT. For SKA, the statistical accuracy in  $P_\theta$  is about 3% around the second and the third peaks (for bin size  $\Delta k = 0.01 h/\text{Mpc}$ ), slightly larger than that of the  $P_g^s$  measurement in the combined sample of SDSS main galaxies and luminous red galaxies

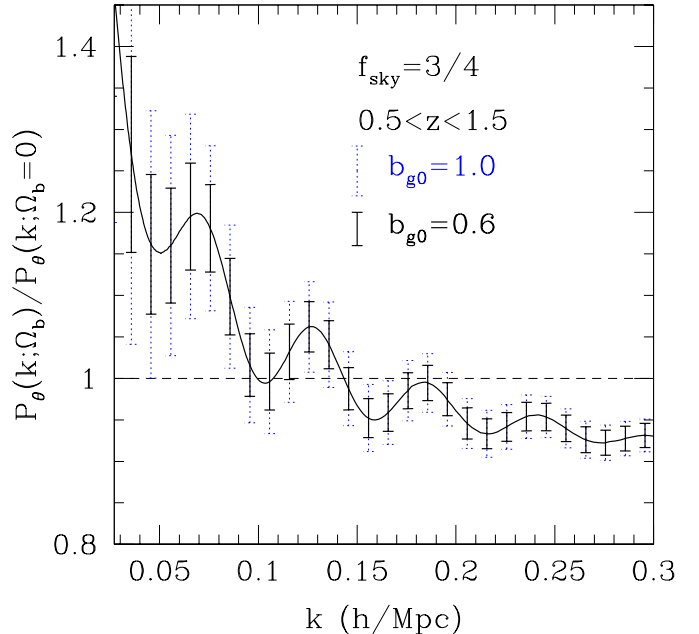


FIG. 3.— Baryon wiggles in  $P_\theta$  of  $0.5 < z < 1.5$  galaxies. The accuracy of  $P_\theta$  reconstruction is sensitive to  $b_g$ . Smaller  $b_g$  causes stronger directional dependence in  $P_g^s$  and weaker statistical fluctuations in  $P_g$  and  $P_{g\theta}$  and thus improves  $P_\theta$  reconstruction. SKA is able to detect numerous galaxies with weak clustering at  $z \sim 1$ .  $P_\theta$  reconstructed through those galaxies can be sufficiently accurate for BAO measurement in the velocity field.

(Percival et al. 2007). SKA extends much larger redshift range than SDSS. Higher redshift means more  $\mathbf{k}$  modes suitable for BAO detection (less severe smear of BAO features by mode-coupling in nonlinear regime) and higher sensitivity to dark matter and dark energy. Thus BAO cosmology through the recovered  $P_\theta$  may gain over current SDSS BAO measurement through  $P_g^s$ .

Accuracy in  $P_\theta$  reconstruction is sensitive to the galaxy bias  $b_g$ . A smaller bias causes a stronger directional dependence and enables better  $P_\theta$  reconstruction. At the first three peaks of BAO, shot noise is in general subdominant for SKA. This allows us the luxury to select those low mass, weakly clustered galaxies for better BAO measurement in the velocity field. A practical question is that, how strong are SKA galaxies clustered. Observationally, we know that galaxies in local universe detected through 21cm line emission are weakly clustered (Meyer et al. 2007). However, we are lack of observations on  $b_g$  of neutral hydrogen mass selected galaxies other than in the local universe. For this reason, we just present a rough estimation here.

SKA galaxies are selected by their neutral hydrogen mass  $M_{\text{HI}}$ . We update the calculation of lower  $M_{\text{HI}}$  limit in Zhang & Pen (2006) with new SKA specifications. In a ten year survey,  $M_{\text{HI}}$  at  $3\sigma$  selection threshold is shown in Fig.2. To calculate the associated bias of these galaxies, we need to convert  $M_{\text{HI}}$  to the galaxy total mass  $M$ . We define the neutral hydrogen fraction  $f_{\text{HI}} \equiv M_{\text{HI}}/M$ . The value of  $f_{\text{HI}}$  is poorly known. But we are able to derive its lower limit. The total neutral hydrogen density is  $\Omega_{\text{HI}} \simeq 3.5 \times 10^{-4}$  today (Zwaan et al. 2005) and in-

creases by a factor of  $\sim 5$  toward  $z \sim 2-3$  (Péroux et al. 2003). Since most baryons and dark matter are not in galaxies while most neutral hydrogen atoms are in galaxies, we expect that  $f_{\text{HI}} \gg \Omega_{\text{HI}}/\Omega_m$ . For several choices of  $f_{\text{HI}}$ , we plot the corresponding  $b_g$  in Fig. 2. We adopt the  $b_g$ - $M$  relation from Jing (1998), whose fitting formulae improves over previous analytical results, especially for low mass halos. Many of SKA detected galaxies at  $z = 1$  are likely well below  $10^{11} M_\odot$ . These galaxies are indeed weakly clustered, with bias as low as  $b_g \sim 0.6$ . These galaxies are very likely numerous, if the shape of neutral hydrogen mass function does not change at low mass end, from local universe to  $z \sim 1$ . Selecting only these galaxies then does not result in significant increase in shot noise. What we gain is a factor of  $\sim 2$  improvement in  $P_\theta$  reconstruction and comparable improvement in BAO measurement, comparing to the case of  $b_{g0} = 1$  (Fig. 3).

## 5. DISCUSSIONS AND SUMMARY

Up to now we only discussed the feasibility to perform self calibration of galaxy bias and to measure BAO in the velocity power spectrum reconstructed from redshift distortion of galaxies. This technique can be applied to any tracers of the large scale structure with precision redshift information. This includes the pseudo 21cm background composed of unresolved 21cm emitting galaxies at  $z \sim 1$  (Chang et al. 2007) and the 21cm background at higher redshifts (Mao & Wu 2007; Wyithe et al. 2007).

It is also feasible to measure BAOs in the velocity field reconstructed from redshift distortion. A fundamental obstacle must be overcome is the statistical fluctuations associated with the reconstruction technique. Since statistical fluctuations in  $P_g$  and  $P_{g\theta}$  propagate into  $P_\theta$  reconstruction, statistical errors in  $P_\theta$  reconstructed are a

factor of  $\sim 10$  times larger than its cosmic variance (Fig. 1). To reach the cosmic variance limit of  $P_\theta$  and improve significantly on BAO measurements in velocity field, new techniques of velocity measurement are needed. Kinetic Sunyaev Zel'dovich effect of galaxy clusters is a possibility. It is also possible to do precision peculiar velocity measurement through millions of SNe Ia with spectroscopic redshifts at  $z \sim 0.5$  (Zhang & Chen 2007). BAO measurement brings extra scientific benefit for such surveys.

$P_\theta$  reconstructed through the galaxy spectroscopic surveys contain valuable information of cosmology, besides the information of the expansion of the universe contained in the BAO. Combining the BAO measurements through galaxy spectroscopic surveys and  $P_\theta$ , we are able to infer the initial fluctuations of the universe, the structure growth of the universe and the geometry of the universe simultaneously. This potentially very promising and powerful application will be addressed in a companion paper.  $P_\theta$  and  $P_{g\theta}$ , when combined with gravitational lensing, also have important applications in probing the nature of gravity at cosmological scales (Zhang et al. 2007, 2008). Finally we emphasize that all these applications rely on precision modeling of redshift distortion. Tremendous work is required to meet the requirement of precision cosmology.

*Acknowledgments:* We thank Ravi Sheth for useful conversations and Ue-Li Pen for useful comments. This work is supported by the one-hundred-talents program of the Chinese academy of science (CAS), the national science foundation of China grant 10533030, the CAS grant KJCX3-SYW-N2 and the 973 program grant No. 2007CB815401.

## REFERENCES

- Abdalla, F. B., & Rawlings, S. 2005, MNRAS, 360, 27  
The Dark Energy Task Force Final Report. Andreas Albrecht, et al. arXiv:astro-ph/0609591  
Angulo, R. E., Baugh, C. M., Frenk, C. S., & Lacey, C. G. 2008, MNRAS, 383, 755. [astro-ph/0702543]  
Bassett, B., et al. 2005, arXiv:astro-ph/0510272  
Bernardeau, F., Colombi, S., Gaztanaga, E. and Scoccimarro, R. 2002, Physics Report, 367, 1-248. [astro-ph/0112551]  
Blake, C., & Glazebrook, K. 2003, Astrophys. J., 594, 665  
Blake, C., Parkinson, D., Bassett, B., Glazebrook, K., Kunz, M., & Nichol, R. C. 2006, MNRAS, 365, 255. [arXiv:astro-ph/0510239]  
Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2007, ArXiv e-prints, 709, arXiv:0709.3672  
Cole, S., et al. 2005, MNRAS, 362, 505  
Crocco, M., & Scoccimarro, R. 2006, Phys. Rev. D, 73, 063519. [arXiv:astro-ph/0509418]  
Crocco, M., & Scoccimarro, R. 2007, ArXiv e-prints, 704, arXiv:0704.2783  
Dekel, A., & Lahav, O. 1999, Astrophys. J., 520, 24  
Eisenstein, D. J., & Hu, W. 1998, Astrophys. J., 496, 605  
Eisenstein, D. J., et al. 2005, Astrophys. J., 633, 560  
Eisenstein, D. J., Seo, H.-J., & White, M. 2007, Astrophys. J., 664, 660  
Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, Astrophys. J., 664, 675  
Glazebrook, K., et al. 2007, arXiv:astro-ph/0701876  
Hu, W., & Sugiyama, N. 1996, Astrophys. J., 471, 542. [astro-ph/9510117]  
Hu, W., & Haiman, Z. 2003, Phys. Rev. D, 68, 063004  
Huff, E., Schulz, A. E., White, M., Schlegel, D. J., & Warren, M. S. 2007, Astroparticle Physics, 26, 351  
Hui, L., Gaztanaga, E., & Loverde, M. 2007, Phys. Rev. D, 76, 103502. [arXiv:0706.1071]  
Hui, L., Gaztanaga, E., & LoVerde, M. 2007, ArXiv e-prints, 710, arXiv:0710.4191  
Jing, Y. P. 1998, ApJ, 503, L9  
Guzik, J., Bernstein, G., & Smith, R. E. 2007, MNRAS, 375, 1329  
Kang, X., Jing, Y. P., Mo, H. J., Börner, G. 2002, MNRAS, 336, 892  
LoVerde, M., Hui, L., & Gaztanaga, E. 2007, ArXiv e-prints, 708, arXiv:0708.0031  
Mao, X.-C., & Wu, X.-P. 2007, ArXiv e-prints, 709, arXiv:0709.3871  
Matarrese, S., & Pietroni, M. 2007, ArXiv Astrophysics e-prints, arXiv:astro-ph/0702653  
Matarrese, S., & Pietroni, M. 2007, ArXiv Astrophysics e-prints, arXiv:astro-ph/0703563  
Matsubara, T. 2007, arXiv:0711.2521  
McDonald, P. 2007, Phys. Rev. D, 75, 043514  
McDonald, P., & Eisenstein, D. J. 2007, Phys. Rev. D, 76, 063009  
Meyer, M. J., Zwaan, M. A., Webster, R. L., Brown, M. J. I., & Staveley-Smith, L. 2007, Astrophys. J., 654, 702  
Pen, U.-L. 1998, Astrophys. J., 504, 601. [arXiv:astro-ph/9711180]  
Percival, W. J., Cole, S., Eisenstein, D. J., Nichol, R. C., Peacock, J. A., Pope, A. C., & Szalay, A. S. 2007, MNRAS, 381, 1053  
Perlmutter, S., et al. 1999, Astrophys. J., 517, 565  
Péroux, C., McMahon, R. G., Storrie-Lombardi, L. J., & Irwin, M. J. 2003, MNRAS, 346, 1103  
Peterson, J. B., Bandura, K., & Pen, U. L. 2006, ArXiv Astrophysics e-prints, arXiv:astro-ph/0606104  
Riess, A. G., et al. 1998, AJ, 116, 1009  
Scoccimarro, R. 2004, Phys. Rev. D, 70, 083007  
Seljak, U. 2001, MNRAS, 325, 1359

- Seo, H.-J., & Eisenstein, D. J. 2003, *Astrophys. J.* , 598, 720
- Seo, H.-J., & Eisenstein, D. J. 2005, *Astrophys. J.* , 633, 575.  
arXiv:astro-ph/0507338
- Seo, H.-J., & Eisenstein, D. J. 2007, *Astrophys. J.* , 665, 14.  
arXiv:astro-ph/0701079
- Sheth, R. K., & Tormen, G. 1999, *MNRAS*, 308, 119
- Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, *Phys. Rev. D*  
, 75, 063512
- Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, *ArXiv*  
*Astrophysics e-prints*, arXiv:astro-ph/0703620
- Spiegel, D. N., et al. 2007, *ApJS*, 170, 377
- Tegmark, M., Hamilton, A. J. S., & Xu, Y. 2002, *MNRAS*, 335,  
887
- Tegmark, M., et al. 2004, *Astrophys. J.* , 606, 702
- Tinker, J. L., Weinberg, D. H., & Zheng, Z. 2006, *MNRAS*, 368,  
85
- Tinker, J. L. 2007, *MNRAS*, 374, 477
- Vallinotto, A., Dodelson, S., Schimd, C., & Uzan, J.-P. 2007, *Phys.*  
*Rev. D* , 75, 103509
- White, M. 2001, *MNRAS*, 321, 1
- White, M. 2005, *Astroparticle Physics*, 24, 334
- Wyithe, S., Loeb, A., & Geil, P. 2007, *ArXiv e-prints*, 709,  
arXiv:0709.2955
- Zhan, Hu., Wang, Lifang, Ponto, P., Tyson, J. A. 2008,  
arXiv:0801.3659
- Zhang, P., & Pen, U.-L. 2006, *MNRAS*, 367, 169
- Zhang, P., Liguori, M., Bean, R., & Dodelson, S. 2007, *Physical*  
*Review Letters*, 99, 141302
- Zhang, P., & Chen, X. 2007, *ArXiv e-prints*, 710, arXiv:0710.1486
- Zhang, P. et al. 2008. in preparation
- Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., & Webster, R. L.  
2005, *MNRAS*, 359, L30